Shape and Motion Categorization for Content-Based Image and Video Database Search

Annual Progress Report

ONR Grant Number: N00014-96-1-0661

Institution: Boston University

PI: Stan Sclaroff

Reporting Period: September 1, 1996 – August 31, 1997

1 Administrative Information

Grant Title:	Shape and Motion Categorization for Content-Based Image and Video Database Search
Grant Number:	N00014-96-1-0661
Institution:	Boston University
PI Name:	Stan Sclaroff
PI Phone:	(617) 353-8928
PI FAX:	(617) 353-6457
PI E-mail:	sclaroffbu.edu
PI WWW Page:	http://www.cs.bu.edu/faculty/sclaroff/
Report WWW Page:	http://www.cs.bu.edu/faculty/sclaroff/onr/
Reporting Period:	Sept 1, 1996 – August 31, 1997
Team Members:	Dr. Stan Sclaroff, Assistant Professor John Isidoro, graduate research assistant, partially-supported by this contract Marco La Cascia, visiting Ph.D. student, partially-supported by this contract Lifeng Liu, graduate research assistant, supported by this contract Leonid Taycher, graduate research assistant, partially-supported by this contract Jonathan Alon, graduate research assistant, to be supported by this contract starting 9/1/97.

Numerical Productivity Measures

	This Period	Contract Total	Career Total
Refereed papers submitted but not yet published:	4	4	4
Refereed papers published:	3	4	28
Unrefereed reports and articles:	0	0	6
Books or parts thereof submitted but not yet published:	1	1	1
Books or parts thereof published:	0	0	5
Patents filed but not yet granted:	0	0	0
Patents granted:	1	1	1
Invited Presentations:	4	6	14
Contributed Presentations:	1	2	26
Honors Received:	0	0	0
Prizes or Awards Received:	0	2	2
Promotions obtained:	0	0	0
Graduate Students Supported:	4	4	5
Post-docs supported:	0	0	0
Minorities supported:	0	0	0

2 Project Summary

The aim of this project is to represent shape categories for interactive, image database search. Rather than directly comparing a candidate shape with all shapes in the database, we will develop methods that describe shapes in terms of their relationship to a few shape prototypes. The underlying representation employs *modal matching,* a deformable shape decomposition that allows users to specify a few example shapes and has the computer efficiently sort the set of objects based on the similarity of their shape. If desired, shapes can be more closely compared in terms of the *types* of nonrigid deformations (differences) that relate them to a few prototype shapes. Furthermore, the original shape can be reconstructed in terms of a linear combination of deformed basis images; thus, a semantics-preserving shape representation will be obtained.

This approach is related to the computer graphics technique of *morphing.* Morphing is accomplished by an artist identifying a large number of corresponding control points in two images, and then incrementally deforming the geometry of the first image so that its control points eventually lie atop the control points of the second image. Using this technique, in-between or novel views can be generated as warps between example views. This suggests an important way to obtain a low-dimensional, parametric description of shape: interpolate between known, prototype views. For instance, given views of the extremes of a motion we can describe the intermediate views as a smooth combination of the extremal views.

All that is required to determine this view-based parameterization of a new shape are: the prototype views, point correspondences between the new shape and the prototype views, and a method of measuring the amount of nonrigid deformation that has occurred between the new shape and each prototype view. The prototypes define a polytope in the space of the (unknown) underlying physical system's parameters. By measuring the amount of deformation between the new shape and extremal views, we locate the new shape in the coordinate system defined by the polytope. This coordinate in prototype space can be used for database indexing and fast search, and for motion tracking and categorization. Such a representation could also prove useful in surveillance (tracking human motion), low bit-rate video compression, target recognition and tracking, and medical image analysis.

This research is built on top of an existing shape representation framework called modal matching. The underlying representation will provide a method for determining point correspondences, warping or morphing one shape into another, and measuring the amount of deformation between an object's shape and prototype views. To achieve the goal of representing shape categories, the modal matching framework will be extended to address four main issues:

Issue 1. Comparison Metrics — To measure a shape's relationship to prototypes, comparison measures will be developed and tested. It is expected that these metrics will fall into two main families: quick metrics that summarize deformation, and detailed metrics that allow closer inspection of how shapes are related.

Issue 2. Category Representation — Automatic methods for selecting the prototypes will be developed. As test databases get larger, work will be done to devise methods for automatically structuring the database into *super categories.*

Issue 3. Including Image Intensity — The formulation will be extended to not only include shape information, but also image (pixel) information. The resulting framework will be used to represent shapes in terms of linear combinations of warped images.

Issue 4. Encoding Motion — The system will be extended to encode rigid, nonrigid, or articulated motion in terms of its similarity to known extremal views. This will enable tracking, describing, and indexing motions in video databases. After sufficient testing, the system will be expanded to include algorithms for figure-ground segmentation.

Summary of Progress During Previous Reporting Period

June 1, 1996 – August 31, 1996. Testing of category representation and distance metrics for sensitivity to noise. In the experiments, the formulation performed significantly better than the moment invariants technique. These results have now been published in *Pattern Recognition*, April 1997.

Summary of Progress During this Reporting Period

The formulation has been extended to include color appearance and lighting changes. *Active Blobs*, a new region-based formulation for nonrigid shape registration, tracking, and recognition is the product of this effort. This new formulation allows us to harness the hardware accelerated triangle texture mapping cabilities becoming prevalent in mid-end workstations and PC's. In tracking experiments, active blobs have been able to achieve a peak rate of five frames/second on an SGI R10K Impact workstation. A paper reporting these results has been accepted as an oral presentation at the International Conference on Computer Vision (see attached technical report).

In addition, preliminary development was started in two areas, resulting in prototype demonstration systems. In the area of moving shape representation and motion-based indexing, we have been testing a preliminary formulation for nonrigid tracking that imbeds the current 2-D active blob in a rigid 3-D surface. This should allow better tracking of convex, rotating/translation 3-D objects. The demonstration application has been head tracking. In the area of nonrigid figure/ground separation, a new formulation for segmentation and detection of deformable shapes is being developed. The approach exploits available information about the statistical distribution of a 2-D shape's deformation and color parameters.

3 Work Plan for Next Year

1. Continue Active Blobs effort. Use active blobs to extend linear combinations of prototypes formulation to include appearance. Target applications: image/video database indexing and nonrigid shape tracking. Formalize and test automatic methods for selecting the prototypes.

2. Continue development of shape tracking algorithms that imbed our 2-D nonrigid shape models in 3-D rigid surfaces. Demonstration application: head tracking. Test on sequences collected under controlled conditions.

3. Continue development of algorithms for model-based figure/ground segmentation.

4 Technical Transitions

1. We share C software developed for this project with Sven Dickinson and Stan Dunn at Rutgers University. The software is being used as part of a pilot project to develop new methods for content-based organization and search for digital image databases of dental X-rays. A paper presenting preliminary results was presented at the IEEE Image and Multidimensional Signal Processing Workshop in March, 1996.

2. The efforts on this project have led to fruitful collaboration with Alex Pentland's group at the MIT Media Lab. Application of results from this project are planned in the area of deformable shape

modeling algorithms for locating and tracking people in dynamic environments. This relationship involves sharing software.

3. The modal matching framework is being independently used and extended by other researchers in Italy and the United Kingdom. Dell'Acqua, Gamba (U. of Pavia, Italy), and Mecocci (U. of Sienna) presented a conference paper reporting the use of modal matching for visual search in image databases using user sketches (in *Proc. International Workshop on Image Databases and Multimedia Search*). For his Ph.D. dissertation work, Mike Syn (Oxford) has extended the modal matching to 3D for use in biomedical dataset analysis.

4. In collaboration with Ron Kikinis at Brigham and Women's Hospital, work is being conducted to transfer deformable shape methods to biomedical applications. The focus is on developing 3-D shape models for tracking and anatomical structures in medical volume data for computer-assisted diagnosis and surgical planning.

5 Significant Accomplishments

- Patent granted for modal matching. U.S. patent 5590261, *A Finite Element for Image Morphing and Alignment*, December 31, 1996. This patent was developed in conjunction with Alex Pentland at the MIT Media Lab.

- New formulation *Active Blobs* developed (report attached). A paper reporting this new formulation was accepted for oral presentation at the International Conference on Computer Vision. This is a very competetive conference, with under 7% of submissions accepted for oral presentation. This work represents continued progress towards addressing the inclusion of image intensity in the models. The formulation can also model changes in lighting.

- Demo system for new nonrigid tracking formulation that imbeds the 2-D nonrigid shape models in a rigid 3-D surface. This may allow better tracking of rotating 3-D objects. The demonstration application has been head tracking.

- Demo system for segmentation and detection of deformable shapes in color images given a known 2-D shape and the statistical distribution of its shape and color parameters. This work may make significant progress towards addressing figure ground segmentation in our system.

6 Publications

Publications Resulting from Work Done on ONR-Managed Grants

1. Sclaroff, S. and Isidoro, J. "Active Blobs," *Proc. International Conference on Computer Vision*, January 1998 (to appear).

2. Sclaroff, S., "Distance to Deformable Prototypes: Encoding Shape Categories for Efficient Search," chapter in *Image Databases and Multi-Media Search*, A.W.M. Smeulders and R. Jain, Ed., World Scientific, Singapore, 1997 (to appear).

3. Sclaroff, S., "Deformable Prototypes for Encoding Shape Categories in Image Databases," *Pattern Recognition*, 30(4):627–642, April, 1997.

4. Sclaroff, S., "Encoding Deformable Shape and Motion Categories for Efficient Content-Based Search," *Proc. First International Workshop on Image Databases and Multimedia Search*, Amsterdam, August 1996.

Publications in Refereed Journals

1. Sclaroff, S., Taycher, L., and M. La Cascia, "ImageRover: A Content-Based Image Browser for the World Wide Web," *Computer Vision and Image Understanding*, special issue on computer vision applications for network-centric computing, August, 1998 (submitted).
2. Martin, J., Pentland, A., Sclaroff, S., and Kikinis, R., "Characterization of Neuropathological Shape Deformations," *IEEE Trans. Pattern Analysis and Machine Intelligence*, (in press).
3. Pentland, A., Picard, R., and Sclaroff, S., "Photobook: Tools for Content-Based Manipulation of Image Databases," *International Journal of Computer Vision*, 18(3), pp. 233–254, June 1996.
4. Sclaroff, S., and Pentland, A., "Modal Matching for Correspondence and Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence* 17(6), pp. 545–561, 1995.
5. Essa, I., Sclaroff, S., and Pentland, A., "A Unified Approach for Physical and Geometric Modeling for Graphics and Animation," *Computer Graphics Forum* 11(3), pp. 129–138, 1992.
6. Pentland, A. and Sclaroff, S., "Closed-Form Solutions For Physically Based Shape Modeling and Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence* 13(7), pp. 715-730, 1991.
7. Sclaroff, S., and Pentland, A., "Generalized Implicit Functions for Computer Graphics," *ACM Computer Graphics*, 25(4), pp. 247–250, 1991.
8. Pentland, A., Essa, I., Friedmann, M., Horowitz, B., Sclaroff, S., and Starner, T., "The Thingworld Modeling System: Virtual Sculpting by Modal Forces," *ACM Computer Graphics* 24(2), pp. 143–144, 1990.

Publications in Refereed Conference Proceedings

1. Taycher, L., La Cascia, M., and Sclaroff, S., "Image Digestion and Relevance Feedback in the ImageRover WWW Search Engine," *Proc. Visual 97*, December 1997 (to appear).
2. Sclaroff, S., Tacyher, L., La Cascia, M., "ImageRover: A Content-Based Image Browser for the World Wide Web," *Proc. IEEE Workshop on Content-Based Retrieval in Image Databases*, June 1997.
3. Zhang, W., Dickinson, S., Sclaroff, S., Marsic, I., Hawkins, S., Feldman, J., Dunn, S., "Searching Medical Image Databases by Image Content," to appear in *Proc. Ninth Image and Multidimensional Signal Processing Workshop*, Belize, March, 1996.
4. Essa, I., Darrell, T., Azarbayejani, A., Sclaroff, S., and Pentland, A., "Looking at People: Extracting Human Movement," *Proc. International Workshop on Computer Vision and Parallel Processing*, Pakistan, January 1995.
5. Sclaroff, S., and Pentland, A., "Physically-based Combinations of Views: Representing Rigid and Nonrigid Motion," *Proc. IEEE Workshop on Nonrigid and Articulate Motion*, Austin, TX, November 1994.

6. Pentland, A., Essa, I., Darrell, T., Azarbayejani, A., and Sclaroff, S., "Visually Guided Interaction and Animation," *Proc. Twenty-Eighth Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, October 1994.

7. Sclaroff, S., and Pentland, A., "Search by Shape Examples: Modeling Nonrigid Deformation," *Proc. Twenty-Eighth Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, October 1994.

8. Ponce, J., Bajcsy, R., Metaxas, D., Binford, T., Forsyth, D., Hebert, M., Ikeuchi, K., Kak, A., Shapiro, L., Sclaroff, S., Pentland, S., and Stockman, G., "Object Representation for Object Recognition," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Seattle, WA, June 1994.

9. Sclaroff, S., and Pentland, A., "On Modal Modeling for Medical Data: Underconstrained Shape Description and Data Compression," *Proc. IEEE Workshop on Biomedical Image Analysis*, Seattle, WA, June 1994.

10. Sclaroff, S., and Pentland, A., "Modal Shape Comparison," *Proc. Workshop on Visual Form*, Capri, Italy, May 1994.

11. Pentland, A., Darrell, T, Essa, I., Azarbayejani, A., and Sclaroff, S., "Visually Guided Animation," *Proc. Computer Animation '94*, Geneva, Switzerland, May 1994.

12. Sclaroff, S., and Pentland, A., "Object Recognition and Categorization Using Modal Matching," *Proc. IEEE CAD-Based Vision Workshop*, Seven Springs, PA, February 1994,

13. Pentland, A, Picard, R., Sclaroff, S., "Photobook: Tools for Content-Based Manipulation of Image Databases," *SPIE Conf. on Storage and Retrieval of Image and Video Databases II*, (SPIE 2185-05), San Jose, CA, February, 1994.

14. Pentland, A., Darrell, T., Azarbayejani, A., and Sclaroff, S., "Towards Machine Vision in Complex Environments," *Proc. Ninth Conf. on Object Recognition and Artificial Intelligence*, Paris, France, January 1994.

15. Sclaroff, S., and Pentland, A., "A Modal Framework for Correspondence and Description," *Proc. Fourth International Conf. on Computer Vision*, Berlin, Germany, May 1993.

16. Sclaroff, S., and Pentland, A., "Modal models: Energy-Based Implicit Functions," *Proc. SPIE Sensor Fusion V*, Boston, MA, November 1992.

17. Sclaroff, S., Essa, I., and Pentland, A., "Vision-Based Animation," *Proc. Eurographics Workshop on Animation and Simulation*, Cambridge, England, September 1992.

18. Pentland, A., Horowitz, B., and Sclaroff, S., "Non-Rigid Motion and Structure from Contour," *Proc. IEEE Workshop on Visual Motion*, Princeton, NJ, October 1991.

19. Sclaroff, S., and Pentland, A., Closed-Form Solutions for "Physically-Based Shape Modeling and Recognition," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Maui, June 1991.

20. Pentland, A., Friedmann, M., Horowitz, B., Sclaroff, S., and Starner, T., "The ThingWorld Modeling System," *Proc. International Workshop on Algorithms and Parallel VLSI Architectures*, pp. 168–172, Pont-a-Mousson, France, June 1990.

21. Darrell, T., Sclaroff, S., and Pentland, A., "Segmentation by Minimal Description," *Proc. Third International Conf. on Computer Vision*, pp. 112–116, Osaka, Japan, December 1990.

Other Major Publications

1. Pentland, A., Picard, R., and Sclaroff, S., "Photobook: Content-Based Manipulation of Image Databases," chapter in *Multimedia Tools and Applications*, B. Furht, Ed., Kluwer International Series in Engineering and Computer Science, Kluwer Academic Publisher, 1996.

2. Pentland, A. and Sclaroff, S., "Modal Representations," chapter in *Object Representation in Computer Vision*, M. Herbert, J. Ponce, T. Boult, and A. Gross, Ed., Lecture Notes in Computer Science series, Springer Verlag, 1995.

3. Sclaroff, S., "Deformable Shape Prototypes for Interactive Image Database Search," abstract in *Proc. NSF/ARPA Visual Information Management Workshop*, Cambridge, MA, June 1995.

4. Pentland, A., and Sclaroff, S., "Modal Representations," abstract in *Report on the NSF/ARPA Workshop on 3-D Object Representation for Computer Vision*, New York, NY, December 5-7, 1994.

5. Pentland, A., Sclaroff, S., Horowitz, B., and Essa, I., "Modal Descriptions for Modeling, Recognition, and Tracking," chapter in *3-D Object Recognition Systems I*, Jain and Flynn, Ed. Elsevier, 1993.

6. Pentland, A., and Sclaroff, S., "From Physics to Phunction," *Proc. Workshop on Functionality*, Harper's Ferry, WV, August 1993.

7. Essa, I., Sclaroff, S., and Pentland, A., "Physically-based Modeling for Graphics and Vision," chapter in *Directions in Geometric Computing*, R. Martin, Ed. Information Geometers, U.K., 1992.

8. Pentland, A., Friedmann, M., Horowitz, B., Sclaroff, S., and Starner, T., "The ThingWorld Modeling System," chapter in *Algorithms and Parallel VLSI Architectures*, E.F. Deprettere, Ed. Elsevier Press, Amsterdam, The Netherlands, 1990.

9. Sclaroff, S., and Pentland, A., "From Features to Solids," abstract in *Proc. AAAI-90 Workshop on Qualitative Vision*, Boston, MA, August 1990.

10. Sclaroff, S., "CSG Ray Tracing Using Octrees," *Proc. Schlumberger Software Conf.*, Ann Arbor, MI, November 1988.

7 Invited Lectures, Panels, Seminars, and Talks:

1. The Rutgers University Series on Human and Computer Vision, Rutgers University: *Active Blobs*, New Brunswick, NJ, August 1997.

2. IEEE Nonrigid and Articulated Motion Workshop: panel on future research directions, Puerto Rico, June 1997.

3. IEEE Workshop on Generic Object Recognition: *Generic Object Recognition with Active Blobs*, Puerto Rico, June 1997.

4. University of Marlyand: *ImageRover: Content-Based Image Browser for the World Wide Web*, College Park, MD, December 1996.

5. Interval Research Corporation: *Deformable Shape Prototypes For Interactive Image Database Search*, Palo Alto, CA, June 1996.

6. Computer Science Colloquium Series, Dartmouth University: *Deformable Shape Prototypes For Interactive Image Database Search*, Hanover, NH, May 1996.

7. The Rutgers University Series on Human and Computer Vision, Rutgers University: *Image Database Search by Example: Modeling Deformable Shape*, New Brunswick, NJ, August 1995.

8. NSF/ARPA Visual Information Management Workshop: panel on Image Databases, Cambridge, MA, June 1995.

9. Institute for Information Technology, National Research Council of Canada: *Deformable Models for Image Understanding*, Ottawa, Ontario, March 1995.

10. Center for Intelligent Machines, McGill University: *Modal Models for Signal Understanding*, Montréal, Québec, March 1995.

11. IEEE Workshop on CAD-based Vision: panel on Object Representation for Computer Vision, Champion, PA, February, 1994.

12. Harvard University Medical School: *Recovering Parametric Physical Models from Medical Images*, Boston, MA, May 1993.

13. Artificial Intelligence Lab, University of Edinburgh: *Modal Models: A Unified Representation for Physics-based Vision and Modeling*, Edinburgh, Scotland, September 1992.

14. INRIA-Rocquencourt: *Modal Analysis for Model Recovery and Recognition, and for Recovering Non-rigid Structure from Motion*, Rocquencourt, France, January 1992.

8 On-Line Information

Web pages describing this project and other research in Boston University's Image and Video Computing Group can be found at: *http://cs-www.bu.edu/groups/ivc/Home.html*. These pages include links to technical reports, project descriptions, team member's home pages, etc.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	19 Aug 97	interim 1 Sep 96 -- 31 Aug 97

4. TITLE AND SUBTITLE
Shape and Motion Categorization for Content-Based Image and Video Database Search: Annual Progress Report

5. FUNDING NUMBERS
G N00014-96-1-0661

6. AUTHOR(S)
Stan Sclaroff

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Computer Science Department
Boston University
111 Cummington Street
Boston, MA 02215

8. PERFORMING ORGANIZATION REPORT NUMBER
sclaroff-ONR-4

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Department of the Navy
Office of Naval Research
Ballston Centre Tower One
800 North Quincy Street
Arlington, VA 22217-5660

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
The aim of this project is to represent shape categories for interactive image database search. The work involves developing representations of images of objects to facilitate content-based retrieval and indexing, especially of deformable objects. This progress report gives detail of work completed from 1 September 1996 to 31 August 1997.

14. SUBJECT TERMS
Content-based image database search, deformable shape, nonrigid motion, indexing

15. NUMBER OF PAGES
~~10~~ 9

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
unclassified	unclassified	unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102